Privacy and Smartphones

Privatics team (Vincent Roca)
Nice University – Sophia Antipolis
February 12th, 2016

Inria Grenoble Rhône-Alpes
Privatics team

contact : vincent.roca@inria.fr

Context: The Mobilitics Inria-CNIL project

- Jan.-2012 – Dec. 2014

- focuses on Android et iOS
  - ❍because they dominate
- analyze personal information leaks in **Apps** and **OS services**
  - ❍compare Android / iOS
  - ❍identify practices et tendencies
  - ❍carry out in-lab and in-vivo studies

Outline

1. introduction
   - ❍ two examples
   - ❍ "personal information" and the French/EU law
2. smartphones and personal information eco-system
   - ❍ why are we here?
   - ❍ let's come back to smartphones
   - ❍ who does what, who earns what?
   - ❍ where is the problem?
3. the manufacturer approaches for privacy parameter control
   - ❍ multiple limits
   - ❍ three quick examples
4. what we learned with the Mobilitics project
   - ❍ a rush towards stable identifiers… for a permanent user tracking that resists to resets
   - ❍ a GPS in your pocket… for the others
5. conclusions

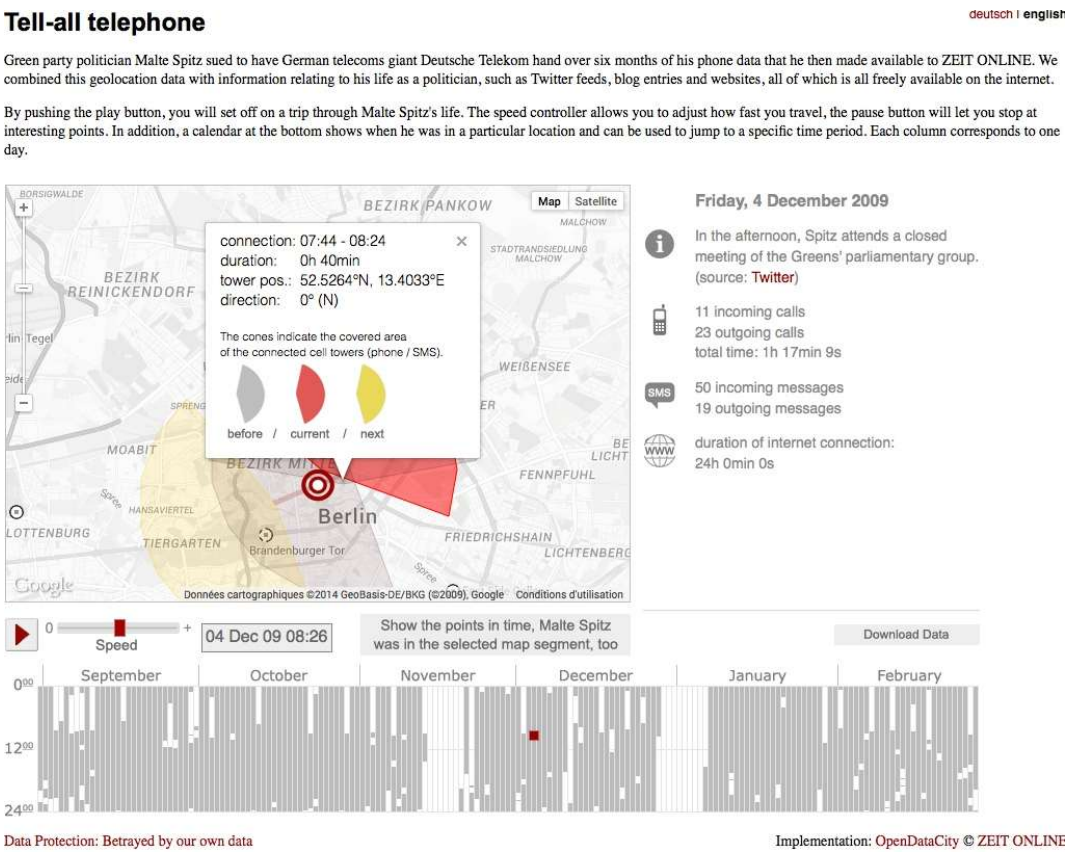
Introduction

- **Two examples to start with…**

Example 1 : geolocation data of a telecom operator (2009)

- Malte Spitz (German Green Party) asked his telecom operator to access his data
  - ❍**enriched with publicly available data (e.g., twitter)**
  - ❍**a dedicated application has been designed to navigate in the history**
    - http://www.zeit.de/datenschutz/malte-spitz-data-retention/

Example 1… (cont')



Example 1… (cont')

- okay, but a legal framework exists that protects the citizens ☺
  - the telecom operator has legal obligations
  - data exists but is only available under specific conditions, after an official request of the authorities

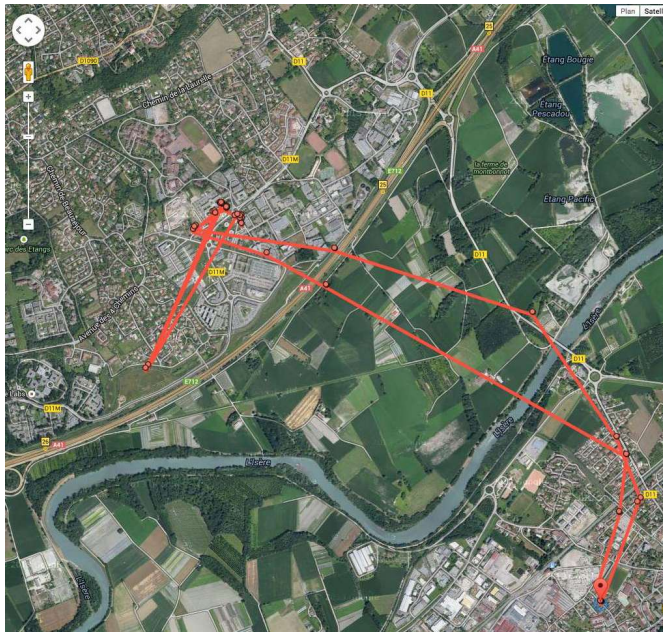
Example 2 : geolocation made in Google

- geolocation collected by my Android smartphone for Google services
  - available
    - NB: login with the gmail account used for the smartphone

    https://maps.google.com/locationhistory/
  - it's worth having a look at it!

NB: Google recently changed this page to hide details! Only a summary is provided. Far less frightening

It is reasonable?



- Google knows where I work, where I live, what I'm doing during the day, how I move…
  - **you too now ;-)**

It is reasonable… (cont.')

- … with an incredible accuracy
  - here is the full list of geolocation points in Google database
    - **a record every 5min during the whole night**
    - **… and every minute during the day if I'm moving!**

« mai 2014 »

lun.	mar.	mer.	jeu.	ven.	sam.	dim.
28	29	30	1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	1
2	3	4	5	6	7	8

Afficher : 1 jour

26 mai 2014

▼ Masquer la date et l'heure

00:00 - 01:00
00:03 00:07 00:12 00:17 00:22 00:26
00:31 00:36 00:41 00:45 00:50 00:55

01:00 - 02:00
01:00 01:04 01:09 01:14 01:19 01:23
01:28 01:33 01:38 01:42 01:47 01:52
01:57

02:00 - 03:00
02:01 02:06 02:11 02:16 02:20 02:25
02:30 02:35 02:39 02:44 02:49 02:54
02:58

03:00 - 04:00
03:03 03:08 03:13 03:17 03:22 03:27
03:32 03:36 03:41 03:46 03:51 03:55

04:00 - 05:00
04:00 04:05 04:10 04:15 04:19 04:24
04:29 04:34 04:38 04:43 04:48 04:53
04:57

05:00 - 06:00
05:02 05:07 05:12 05:16 05:21 05:26
05:31 05:35 05:40 05:45 05:50 05:54
05:59

06:00 - 07:00
06:04 06:09 06:13 06:18 06:23 06:28
06:32 06:37 06:42 06:47 06:51 06:56

07:00 - 08:00
07:01 07:06 07:10 07:15 07:20 07:25
07:29 07:34 07:39 07:44 07:48 07:49
07:50 07:51 07:52 07:53 07:54 07:55
07:56 07:57 07:58 07:59

08:00 - 09:00
08:00 08:01 08:02 08:03 08:04 08:05
08:06 08:07 08:08 08:09 08:11:05
08:11:59 08:12 08:18 08:21 08:24
08:25 08:26 08:27 08:28 08:29 08:30
08:31 08:32 08:37 08:42 08:47 08:51
08:56

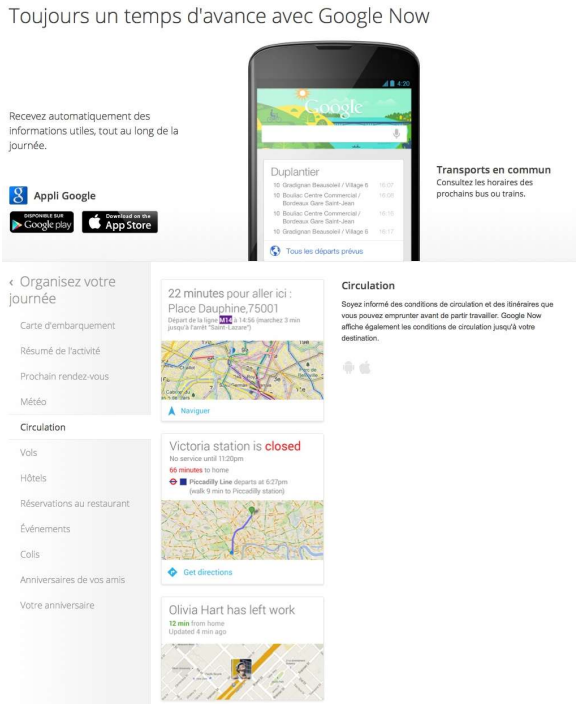
09:00 - 10:00
09:01 09:06 09:10 09:15 09:20 09:25
09:29 09:34 09:39 09:44 09:48 09:53
09:58

10:00 - 11:00
10:03 10:07 10:12 10:17 10:22 10:26
10:31 10:36 10:41 10:45 10:50 10:55

11:00 - 12:00
11:00 11:04 11:09 11:14 11:19 11:23
11:28 11:33 11:38 11:42 11:47 11:52

It is reasonable… (cont.')

- why is it so?
  - I've enabled Google Now : http://www.google.com/landing/now/



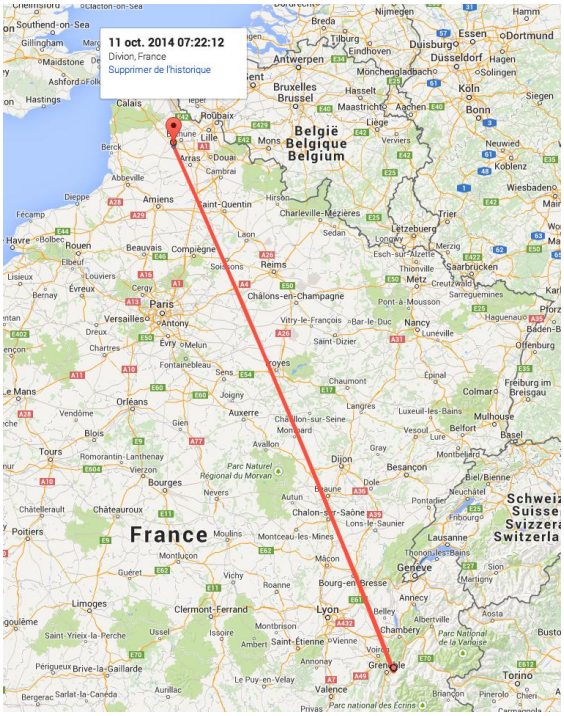
It is reasonable… (cont.')

- of course…
  - Google Now can be disabled (OFF by default)
  - I can reset geolocation data on Google web site

- but…
  - isn't it **disproportionate** with respect to the service provided?
    - **there's a general principle: "collect the minimum needed to provide a given service"**
    - **does the service require to keep all the records in the database for long periods?**
  - there are also geolocation **errors**…

It is reasonable… (cont.')

- at Grenoble at 7:20, in the north 2 minutes later
  - **there the mistake is obvious but sometimes it's credible!**



Introduction

- **"Personal Information" (PI) and the French/EU law**

Loi informatique et liberté (1978)

identity is not required as long as a path to an identity can be found

"Constitue une donnée à caractère personnel **toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement**, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres. Pour déterminer si une personne est identifiable, il convient de **considérer l'ensemble des moyens** en vue de permettre son identification dont dispose ou auxquels peut avoir accès le responsable du traitement **ou toute autre personne**."

no limit on the technical means

no limit: anybody in the world

Loi informatique et liberté (1978)… (cont.')

- the **nature** of the information does not matter…
  - **can be anything (e.g., temperature in a home)**
- …if there is a link to a person, it's a Personal Info (PI)
- this link can be **direct**…
  - **e.g., we record temperature + name**
- or **indirect**
  - **e.g., we record temperature + EDF client ID**
- a person is considered identifiable if the **data controller** has the information to identify him
  - **e.g., EDF collects your home temperature + EDF client ID**
- or **anybody else** in the world
  - **e.g., EDF collects your home temperature + IP address of the sensor. Here the ISP can link the IP to the ADSL user**

Loi informatique et liberté (1978)… (cont.')

- French and EU definition of PI is very broad
  - **in US the linkability to a person is restricted only to the data controller (i.e., database owner)**
  - **MAJOR DIFFERENCE!**

- NB: a common term, PII (Personally Identifiable Information)

Loi informatique et liberté (1978)… (cont.')

- Question 1: what about the following claim?

  "we don't collect your name, age or address, only non personal information"

  - **wrong if linkability to a person remains possible**

- Question 2: is an IP address a PI?

  - **yes in France and in EU**
  - **no in the US, apart from the ISP**

Loi informatique et liberté (1978)… (cont.')

- **sensitive information** CANNOT be collected/ processed

  « Il est **interdit** de collecter ou de traiter des données à caractère personnel qui font apparaître, **directement ou indirectement**, les **origines raciales ou ethniques**, les **opinions politiques, philosophiques ou religieuses** ou **l'appartenance syndicale** des personnes, ou qui sont relatives à la **santé** ou à la **vie sexuelle** de celles-ci. »

- it's clear, non ambiguous: it's prohibited

- in practice it's pretty complex because of inference
  - **if Google knows I'm at a church every Sunday morning (thanks to geolocation) he knows something whose collection is prohibited**

Loi informatique et liberté (1978)… (cont.')

- many obligations to the data controller

  « 1° Les données sont collectées et traitées de manière loyale et licite ;

  fair collection

  2° Elles sont collectées pour des finalités déterminées, explicites et légitimes et ne sont pas traitées ultérieurement de manière incompatible avec ces finalités. […];

  well defined goal

  3° Elles sont adéquates, pertinentes et non excessives au regard des finalités pour lesquelles elles sont collectées et de leurs traitements ultérieurs ;

  collect the bare minimum

  4° Elles sont exactes, complètes et, si nécessaire, mises à jour ;[…] ;

  accuracy

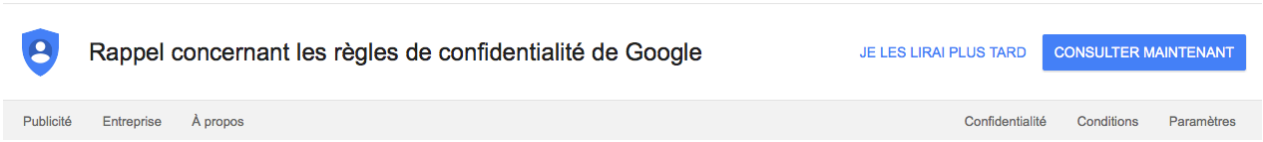
  5° Elles sont conservées sous une forme permettant l'identification des personnes concernées pendant une durée qui n'excède pas la durée nécessaire aux finalités […]. »

  limited duration

Ways to escape the PI rules

the data collector can do a lot if…

- solution 1: they get the **free and informed consent of the user**
  - ❍**"consentement libre et éclairé"**
  - ❍**explains why Google urges the user to read their confidentiality rules**



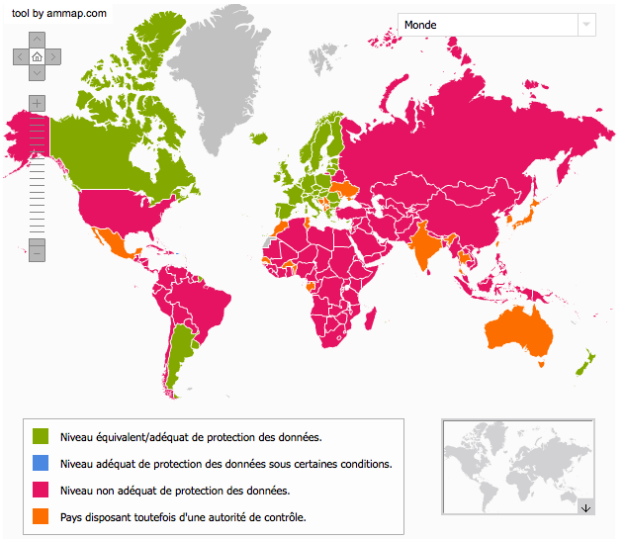
  - ❍**is it sufficient?**
    - no if the user is not free to use the service (no alternative)
    - no if the privacy rules are not compliant with French / EU law (ex. Facebook)

Ways to escape the PI rules… (cont.')

- solution 2: data is **anonymized**
  - ❍**if linkability to a person is impossible it is no longer PI**
  - ❍**but secure anonymization can be pretty hard to achieve**
    - because of inference attacks with side information
  - ❍**and not necessarily sufficient**
    - if a group of people is known to have a certain property, and if I'm known to belong to this group, even if my individual record cannot be identified in the database, one knows I have this property too

PI transmission beyond EU

- personal info cannot be sent beyond EU borders
  - **there are exceptions for countries whose data protection law is compliant with that of EU**



  - **there are exception for companies who signed a specific contract**

PI transmission beyond EU… (cont.')

- close-up on **US companies**
  - US is not recognized as trustworthy W.R.T. PI protection
  - the "Safe Harbor" program was used to authorize PI collection till Oct. 2015
    - **EUJC judgment (Max Schrems) concluded the US law does not guaranty the security of EU citizens PI**
  - no rule today and PI collection is therefore prohibited…
    - **… but negotiations are on the way to establish new legal foundations**
  - in the meantime high pressure of US companies to get the "free and informed" user consent

Outline

1. introduction
   - ❍ two examples
   - ❍ "personal information" and the French/EU law
2. smartphones and personal information eco-system
   - ❍ why are we here?
   - ❍ let's come back to smartphones
   - ❍ who does what, who earns what?
   - ❍ where is the problem?
3. the manufacturer approaches for privacy parameter control
   - ❍ multiple limits
   - ❍ three quick examples
4. what we learned with the Mobilitics project
   - ❍ a rush towards stable identifiers… for a permanent user tracking that resists to resets
   - ❍ a GPS in your pocket… for the others
5. conclusions

● Why are we here?

A massive worldwide surveillance

- we leave traces that are **systematically** recorded whenever we use Internet and our smartphone
  - **on the "visible" web**
  - **on the "invisible" web**
  - for **economic** or **security** reasons

Surveillance on the "visible" web

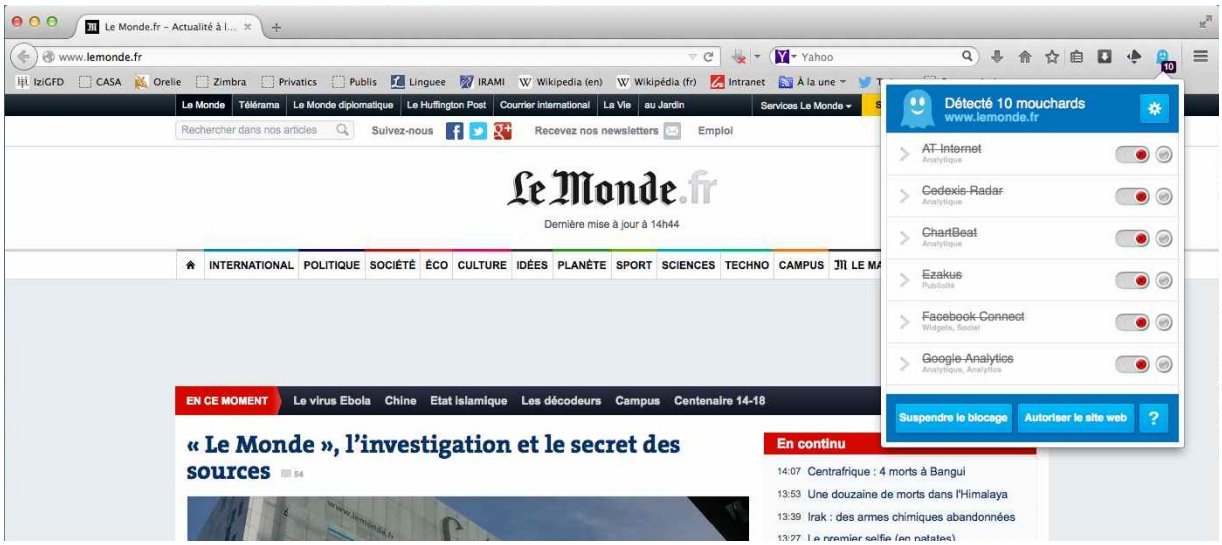
- Foursquare knows **where you are**
- Flickr knows **what you are watching**
- Facebook knows **what you're doing**
- Linkedin knows **where and with whom you're working**
- Twitter knows **what you're saying**
- Amazon knows **what you're buying**
- Google knows **what you're thinking**
- and much more…

If we cross all information, it's becoming terrifying

http://www.le-tigre.net/Marc-L.html

Surveillance on the "invisible" web

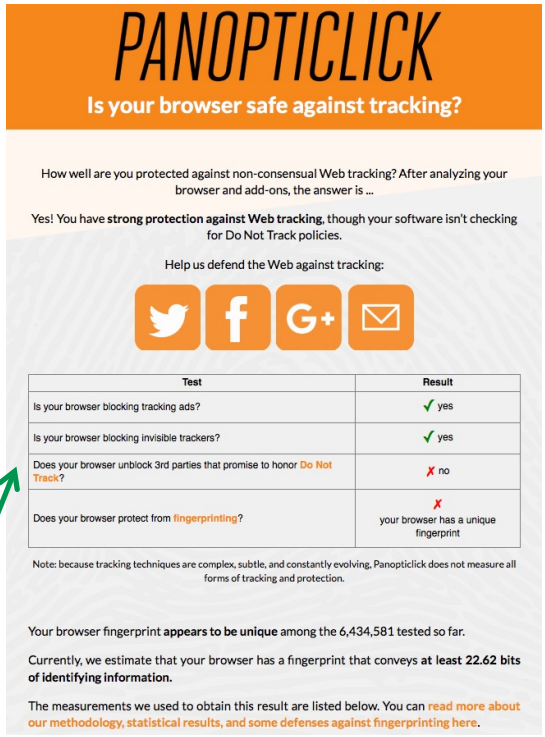
- thanks to cookies, pixels, "I like" buttons, etc. of web sites
  - one can easily **track** and **profile** all users



Surveillance on the "invisible" web

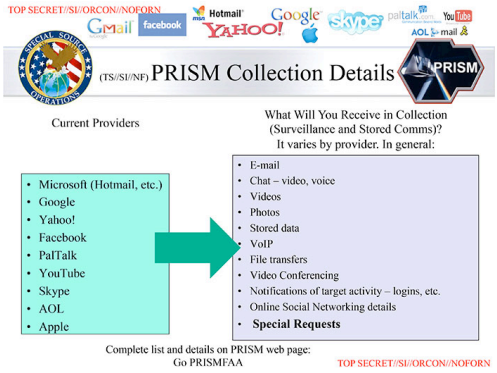
- even if you don't provide your ID, anyway your browser is unique in the world and can be tracked
  - Panopticlick
    - **fingerprinting based on config, version, OS, screen resolution, etc.**
  - add blockers do help but are not 100% efficient

I'm using Adblock, Ghostery and Privacy badger!



This situation can easily lead to abuses

- NSA…
  - the issue is not to track well identified targets, but
    - **to track all the world citizens**
    - **to compromise the security of our tools**



- and NSA is not the only agency that does it

- **let’s come back to smartphones…**

Smartphones have a key responsibility

- our everyday "companions"…
  - useful, always connected, easy to customize
- but they also

concentrate

personal information

when we use them: phone calls, SMS, web, applications, etc.

generate

personal information

GPS, NFC, WiFi, camera, fingerprint sensor, heart rate sensor, etc.

A key responsibility… (cont.')

- they know a lot on our cyber-activities
  - applications generate many **opportunities** to leak personal information
    - **it justifies that web site you visit invite you to download and use their own App…**

"notre mouchard de poche préféré ?"

What is the subject of this talk?

- a smartphone is composed of
  - an application processor
  - an operating system (OS)
    - **Android (Google), iOS (Apple), Windows Phone, FirefoxOS ✝, Tizen, Cyanogen OS, etc.**
  - applications (“Apps”)

our subject (Android/ iOS)

  - a full system (processor + OS) for baseband communications
    - **hidden, no open spec, closed industry**

very complex to study

http://events.ccc.de/congress/2011/Fahrplan/attachments/2022_11-ccc-qcombbdbg.pdf

- **Who does what, who earns what?**

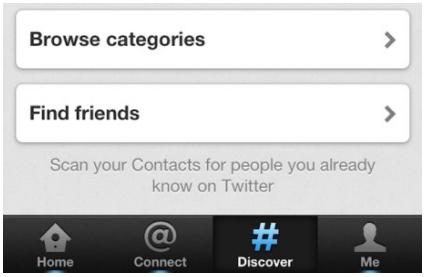
A complex eco-system

- **complex** because several actors are involved
  - « first party » : owns the App
    ⇒ those we see
  - « third party » : Advertising and Analytics (A&A)
    ⇒ those we never see
  - the third party has clients (e.g., advertising companies)
  - certain actors play multiple roles (e.g., Google and Facebook)

- it's impossible to trust everybody
  - two examples…

Example 1: information leaks "by error"

- Twitter (Feb. 2012):
  - "La fonctionnalité de recherche d'amis de […] Twitter permet au service en ligne de télécharger sur ses serveurs les carnets d'adresses et la liste de contacts des utilisateurs. Une fois téléchargées sur ses serveurs, ces données sont conservées 18 mois."



http://www.zdnet.fr/actualites/twitter-copie-et-conserve-18-mois-sans-consentement-les-carnets-d-adresses-des-utilisateurs-39768632.htm

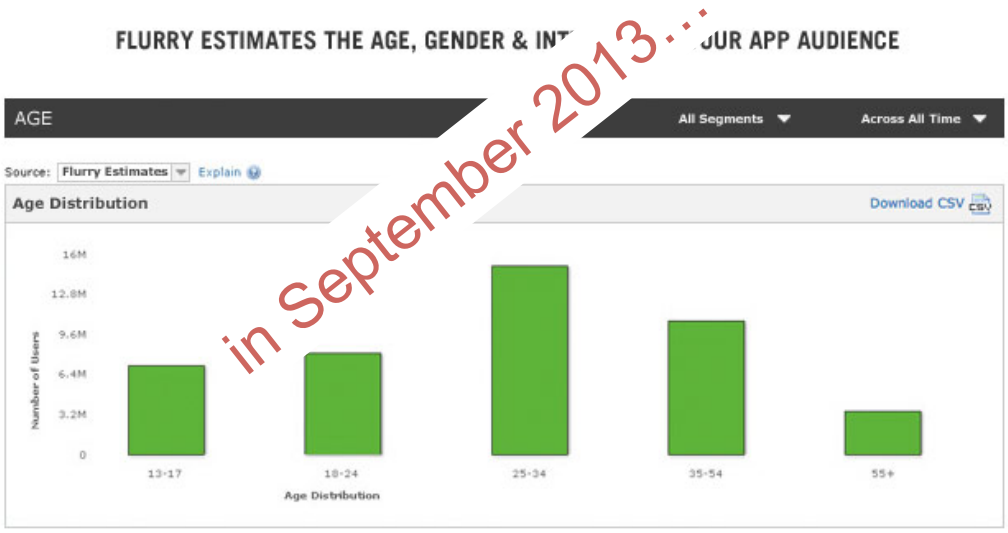
  - similar scandals happened with LinkedIn et Path en 2012!

- those are strategic errors
  - immediately fixed in a new version of the App
  - reputation is essential for those companies and risks are huge

Example 2: massive, organized collection

- Flurry (from Yahoo)
  - http://www.flurry.com

The enormous amount of data Flurry handles directly translates into unique, powerful insights for you. The service takes in over 3.5 billion app session reports per day totaling more than 3 terabytes, and our storage is in the petabytes. Here are some examples of how we use big data to create advantages for you:

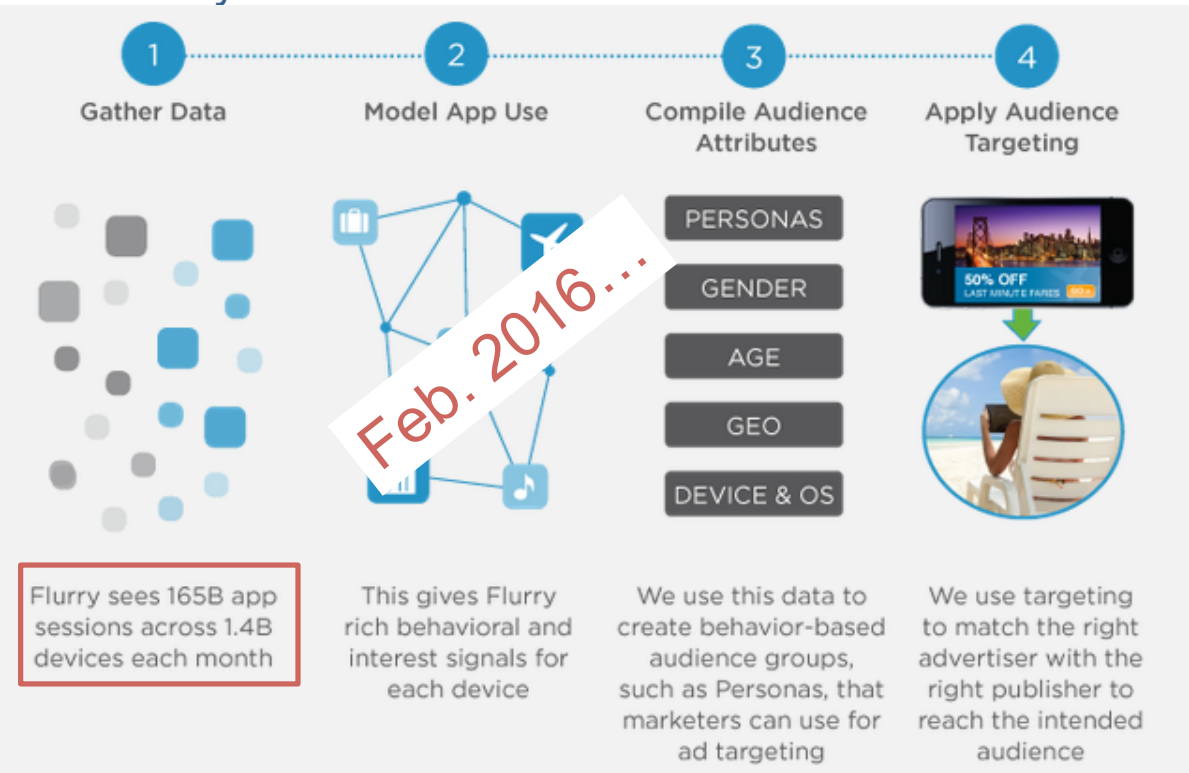


Example 2: massive collection… (cont.')

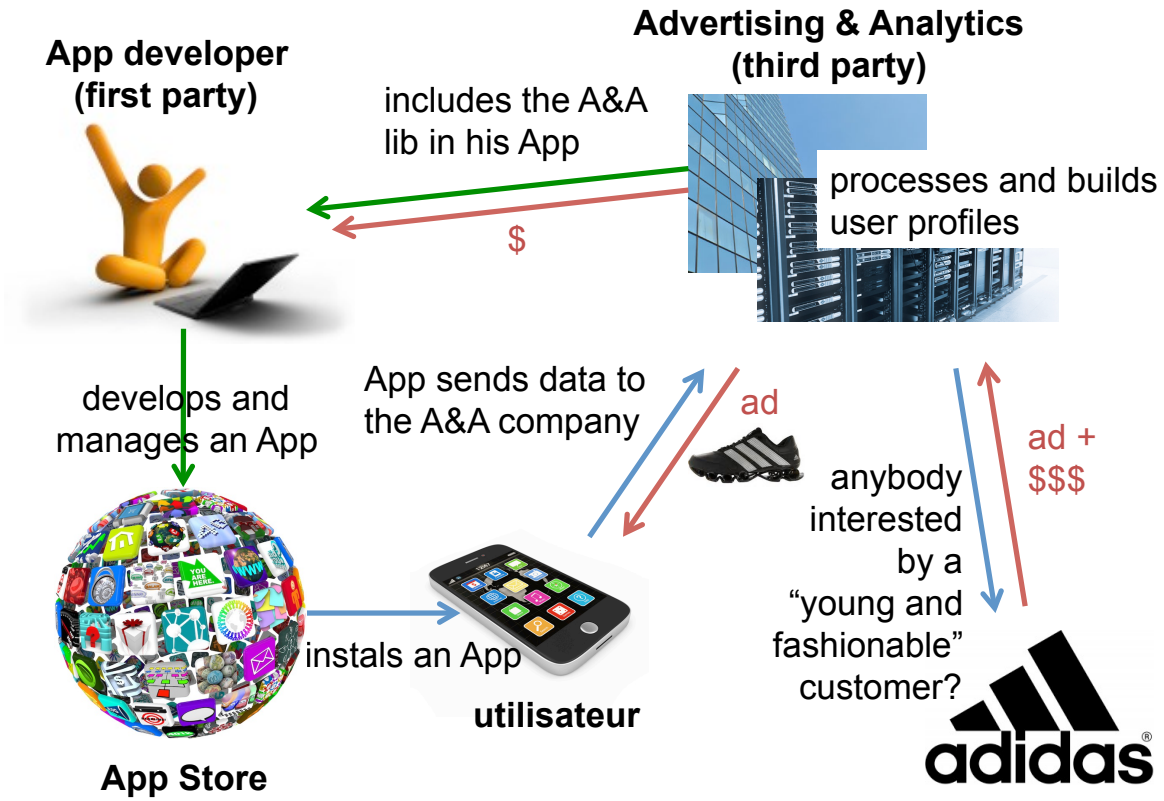
- what for?
  - in order to **track users**
    - **does the same user come back? What Apps does he use? With what frequency? When?**
  - in order to **profile users**
    - **is he a middle-age man? Does he like sport, technology? Does he read news, etc.**
- final goal is to
  - sell **targeted advertising** on the smartphone
    - **high click ratio because ad is targeted**
  - but this database can easily be user for other purposes…
    - **massive surveillance**

Example 2: massive collection… (cont.')

- Ex. Flurry



The actors and their relationships



About mobile advertising

- many companies

AdMob by Google

millennial media

iAd

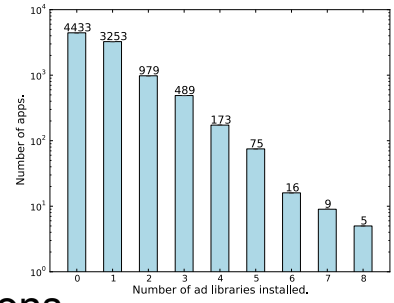
FLURRY

criteo

  - **>8 B$** of revenues for mobile advertising in 2013 for Google

About mobile advertising… (cont.')

- a few facts for Android (2011 data)
  - 77% of 50 free Apps are supported by ad. [1]
  - 35% of free Apps use at least two A&A libraries [2]
    - **in the hope to earn more?**



  - A&A libs require additional authorizations
    - **a free App asks 2-3 additional authorizations WRT paying version of the App [1]**

- **[1] "Don't kill my ads! Balancing Privacy in an Ad-Supported Mobile Application Market", HotMobile 2012.**
- **[2] "AdSplit: Separating smartphone advertising from applications", Usenix Security 2012.**

- **where is the problem?**

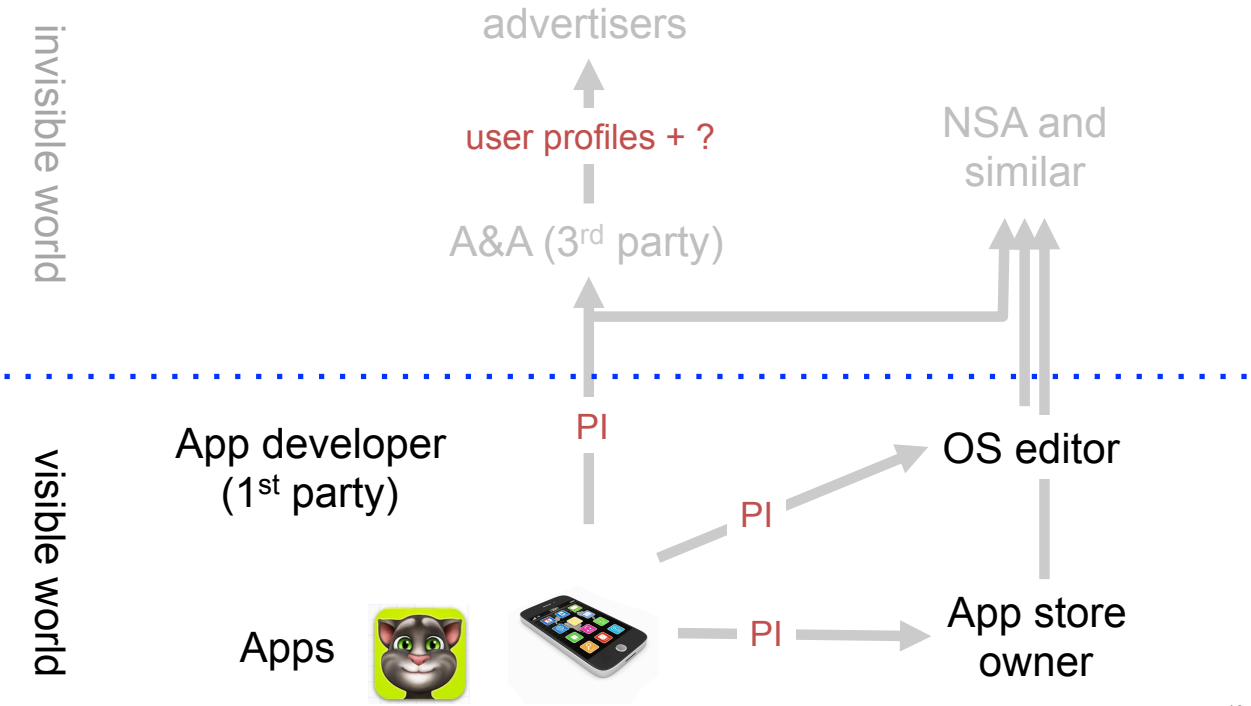
Where is the problem?

- just another business model?

> « Les données personnelles sont le nouveau pétrole de l'internet et la nouvelle monnaie du monde numérique. »
>
> ***M. Kuneva, Commissaire europ. à la consommation, 2009***

- maybe the price to pay for free Apps/services, but…

1- the ecosystem is so complex we cannot trust all actors



2- unreasonable practices

- a collect of our PI that is:

MASSIVE

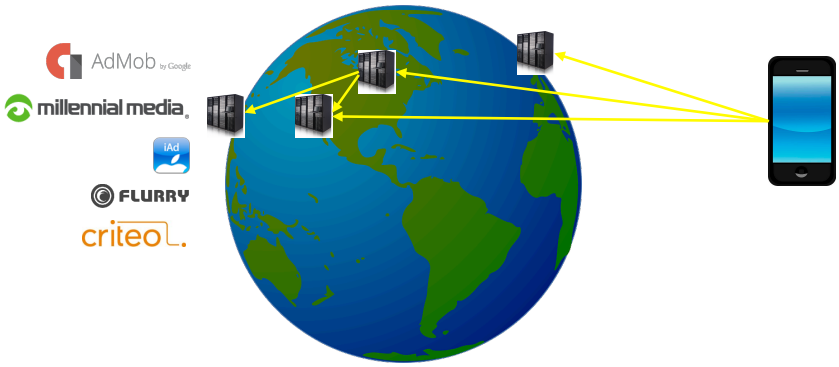
disproportionate

unnoticed

- It's not in line with FR/EU law

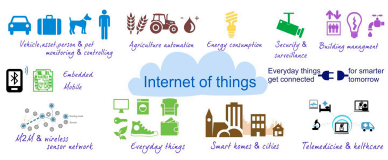
3- uncontrolled collection of our PI

- data is immediately **exfiltrated** beyond EU in order to be stored, processed or exchanged in unknown conditions, **without any control**
  - **FR and EU laws apply difficultly in those countries**
- under FR law, a user must be able to access, correct and withdraw his PI which is not always the case here!



And it's just the begining

- PI collection will be more and more intrusive:
  - generalization of smartphone payment
  - wearable connected devices
  - home connected appliances
    - **e.g., intelligent thermometer**
  - "quantified self" trend
  - connected cars
  - IoT



Outline

1. introduction
   - ❍ two examples
   - ❍ "personal information" and the French/EU law
2. smartphones and personal information eco-system
   - ❍ why are we here?
   - ❍ let's come back to smartphones
   - ❍ who does what, who earns what?
   - ❍ where is the problem?
3. the manufacturer approaches for privacy parameter control
   - ❍ multiple limits
   - ❍ three quick examples
4. what we learned with the Mobilitics project
   - ❍ a rush towards stable identifiers… for a permanent user tracking that resists to resets
   - ❍ a GPS in your pocket… for the others
5. conclusions

● The manufacturer approaches for privacy parameter control

Complementary approaches

- several approaches
  - **market centric:** the market owner checks the App before accepting it

App Store

Google play

  - **user centric:** ask for the user consent…
    - … upon **installing** the App

Google

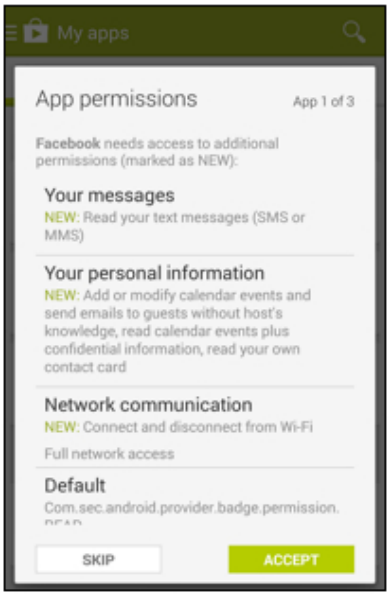
    - … or **dynamically**, when using the App

About installation based authorizations

Reminder: Google/Android

Google

- an App having specific requirements asks for user consent at installation time
  - **responsibility** is transferred to the user
  - very basic approach



About dynamic authorizations

Reminder: Apple/iOS

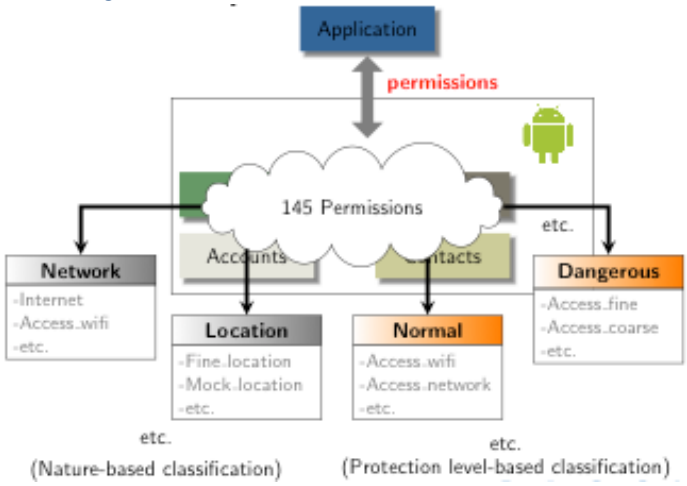
(also quickly introduced in Android 4.3, then removed)

- a dedicated control panel enables users to authorize or ban access to PI of each App
  - **responsibility** is transferred to the user but this latter can change its mind at any time
  - here since iOS 6… and progressively improved

- **many limits for these approaches**

A complex authorization system…

- 145 different types of authorizations



- users won't necessarily understand the **implications**
  - **example : ACCESS_WIFI_STATE**
    - many PI can be inferred without the user being aware of it

…that is also extremely limited

- accept or go elsewhere
  - **we're not living in a binary world!**
- no **behavioral** control of the App
  - **authorizing an App to access my location and Internet for a punctual service does not mean I authorize this App to access my geolocation every minute and to send it to foreign servers**
- no control on the **composition** of authorizations
  - **authorizing an App to access my contacts and Internet does not mean I authorize this App to SEND my contacts to remote servers**

What about Apple?

- much better, but not yet sufficient
- no **behavioral** control of the App
  - **idem**
  - **authorizing access to a PI does not mean I authorize any access and processing modality for this PI**

- **three quick examples**
  - **RATP** App, mid-2013 version
  - an example of **Android authorization** with major implications…
    - J. P. Achara, M. Cunche, V. Roca, A. Francillon, « Short paper: WifiLeaks: Underestimated Privacy Implications of the ACCESS WIFI STATE Android Permission », IEEE WiSec'14.
  - **tracking users thanks to the Wifi interface of their smartphone,** slides from M. Cunche (journées du LERTI)

Ex. 1: don't blindly believe the App Privacy Policies (RATP, version 5.4.1)

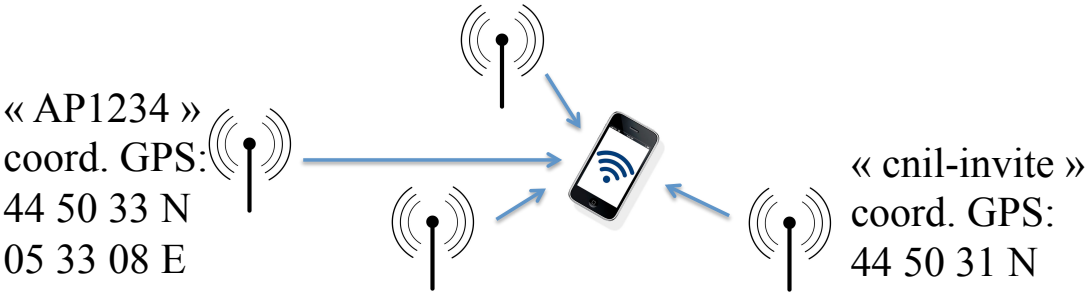
- "there's no problem" according to RATP
- really?
  - the active App list, my MAC address, smartphone name, accurate geolocation, a stable identifier are sent to Adgoji (ssl) and Sofialys (cleartext!)
- see our blog : part-1 et part-2: https://team.inria.fr/privatics/

Ex. 2: an Android authorization with unexpected implications…

- imagine an App, that without asking the user explicit authorization…
- … can **track** the user thanks to a stable identifier
  - it's the Wifi MAC **address**
    - **e.g.** `68:a8:6d:28:ce:1f`
  - guaranteed to be **unique** in the world
  - **impossible** to re-initialize

Ex. 2: an Android authorization with unexpected implications…

- imagine an App, that without asking the user explicit authorization…
- … knows your **location**
  - by listening **Wifi networks in range**, then thanks to a broad database giving the geolocation of all AP can locate the smartphone by triangulation
  - in urban environments, can be **very accurate**



Ex. 2: an Android authorization with unexpected implications…

- imagine an App, that without asking the user explicit authorization…
- … knows a part of your **travels** and your **profile**
  - via the list of Wifi AP to which you connected, which is automatically registered in your smartphone

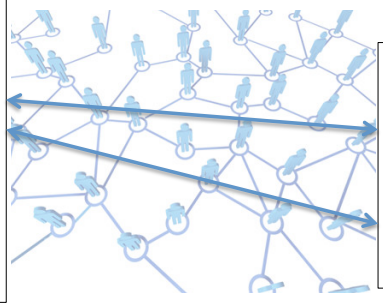
terminal 68:a8:6d:28:ce:1f

```
eduroam
Inria
monwifiamoi
aeroportdelyon
hilton
globecom2014
cnil-invite
```

Ex. 2: an Android authorization with unexpected implications…

- imagine an App, that without asking the user explicit authorization…
- … can infer **social links** between users
  - by calculating the distance between their Wifi connection list, after creating a large dedicated database

terminal 68:a8:6d:28:ce:1f
eduroam
inria
monwifiamoi
aeroportdelyon
hilton
globecom2014
cnil-invite



terminal 68:a8:78:a8:b1:21
McDonalds' free wifi
trucmuchewifi
inria
freewifi
monwifiamoi

Ex. 2: it's possible thanks to Android!

- it is sufficient to ask the **ACCESS_WIFI_STATE** and **INTERNET** authorization at installation time…
  - no user can imagine this is possible
  - and the authorization descriptions gives no clue!

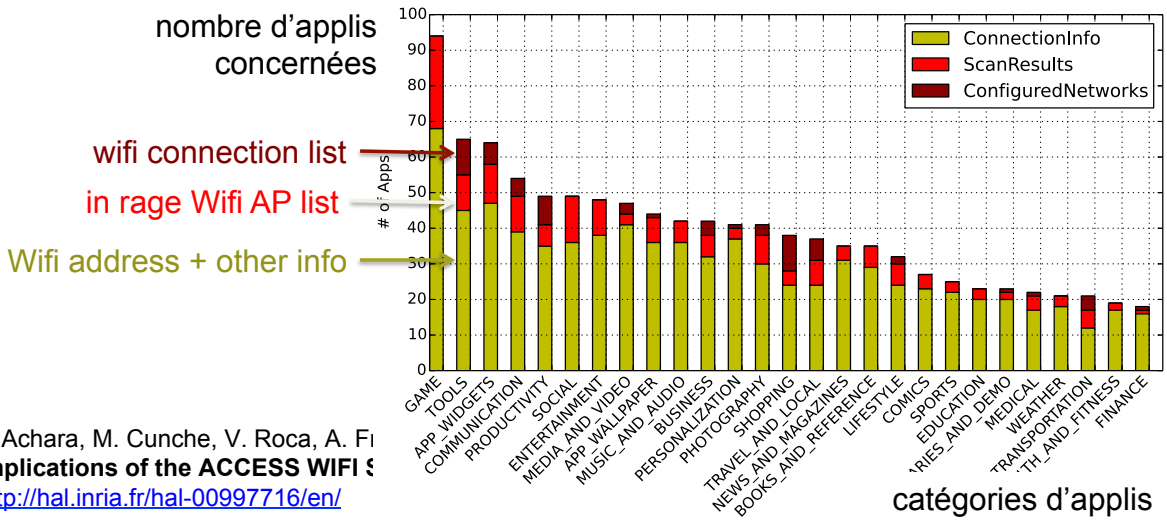
Network communication

View Wi-Fi connections

Allows the app to view information about Wi-Fi networking, such as whether Wi-Fi is enabled and name of connected Wi-Fi devices.

Is it in use?

- **Yes…** Within the 2700 most popular Apps, 41% ask both permissions and many of them use it



J. Achara, M. Cunche, V. Roca, A. Fi
Implications of the ACCESS WIFI S
http://hal.inria.fr/hal-00997716/en/

Ex. 3 : tracking users in physical world thanks to their smartphone Wifi interface

- Wi-Fi tracking system[11]
  - Set of sensors collect Wi-Fi signal
  - Detect and track Wi-Fi devices and their owners
  - MAC address used as identifier

M. Cunche slide (Inria, Privatics)

[11] A. B. M. Musa and Jakob Eriksson. "Tracking unmodified smartphones using Wi-Fi monitors". In: *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*. 2012.

Ex. 3 : tracking users… (cont.')

M. Cunche slide (Inria, Privatics)

- Physical analytics
  - Similar to Web Analytics
  - Frequency and length of visit, number of visitor, peak hour ....
- Trajectory reconstruction
  - Signal received by several sensors
  - Triangulation based on signal strength

Outline

1. introduction
   - ❍ **two examples**
   - ❍ **"personal information" and the French/EU law**
2. smartphones and personal information eco-system
   - ❍ **why are we here?**
   - ❍ **let's come back to smartphones**
   - ❍ **who does what, who earns what?**
   - ❍ **where is the problem?**
3. the manufacturer approaches for privacy parameter control
   - ❍ **multiple limits**
   - ❍ **three quick examples**
4. what we learned with the Mobilitics project
   - ❍ **a rush towards stable identifiers… for a permanent user tracking that resists to resets**
   - ❍ **a GPS in your pocket… for the others**
5. conclusions

- **What we learned with the Mobilitics project**
  - **"in-vivo" experiments, with volunteers using an instrumented iOS or Android smartphone during 3 months in a daily basis**

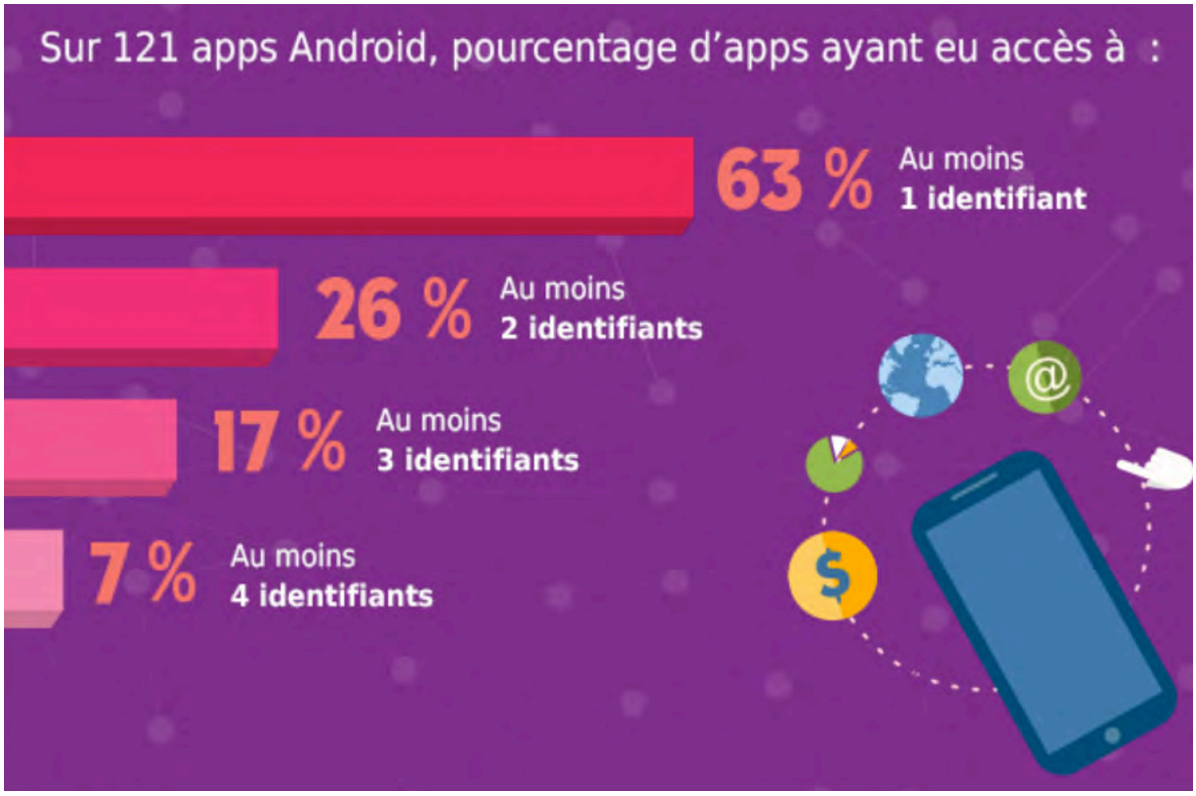
Transparents CNIL lors de la conférence de presse Inria-CNIL du 15/12/2014

CNIL

A rush towards stable identifiers

Résultats généraux, comparaison entre les deux saisons	**iOS 5** (tests de novembre 2012 à janvier 2013)		**Android « Jelly Bean »** (tests de juin à septembre 2014)	
Nombre d'applications	**total : 189**		**total : 121**	
Qui communiquent sur le réseau	176	93%	80	66%
Qui accèdent à l'UDID/android ID	87	46%	41	34%
Qui accèdent à la géolocalisation	58	31%	29	24%
Qui accèdent au carnet d'adresses	15	8%	20	17%
Qui accèdent au calendrier	3	2%	4	3%
Qui accèdent au nom de l'appareil	30	16%	non mesuré	
Qui accèdent au nom d'opérateur	non mesuré		28	23%
Qui accèdent à l'IMEI (identité d'équipement mobile)	non mesuré		24	20%
Qui accèdent à l'adresse MAC WiFi	non mesuré		9	7%
Qui accèdent au numéro de téléphone	non mesuré		7	6%
Qui accèdent à l'identifiant de carte SIM (ICCID)	non mesuré		6	5%
Qui accèdent à la liste des points d'accès WiFi (SSID)	non mesuré		5	4%

A rush towards stable identifiers… (cont.')



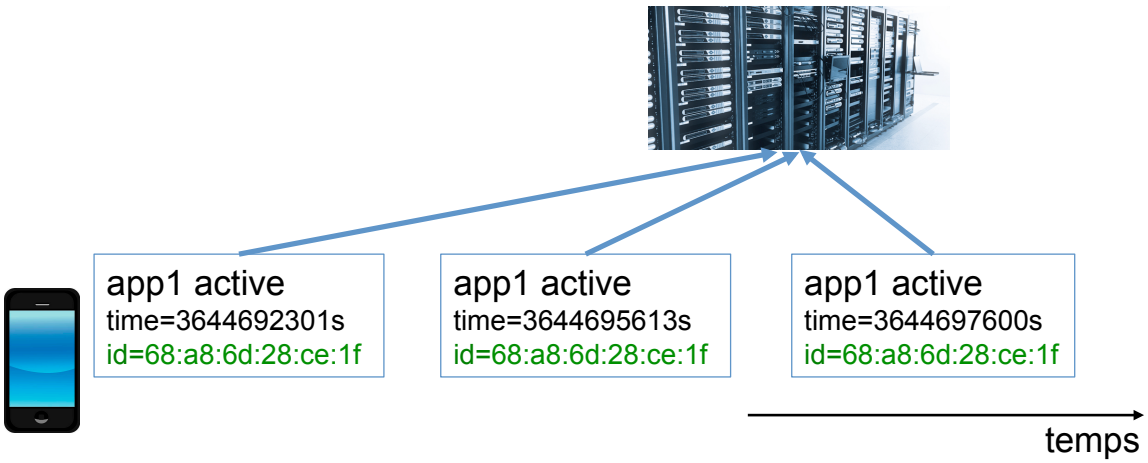
About stable identifiers and their use

- AndroidID
  random number generated upon starting the smartphone for the first time and kept in a stable memory
- MAC address of Wifi (or Bluetooth) interface
  identifies uniquely the network interface (e.g., `68:a8:6d:28:ce:1f`)
- IMEI (International Mobile Equipment Identity)
  uniquely identifies a smartphone (used for instance to block a stolen phone)
- IMSI (International Mobile Subscriber Identity)
  identifies a user at his/her cell phone operator
- AdID (Advertising Identifier)
  special ID used for advertising tracking that a user can reset at any time to prevent long term tracking (in theory at least)

About stable IDs and their use… (cont.')

- looks safe but…
  - ❍**considered as PI by FR/EU law**

- **stable IDs** are perfect for **tracking** users on the long term



About stable IDs and their use… (cont.')

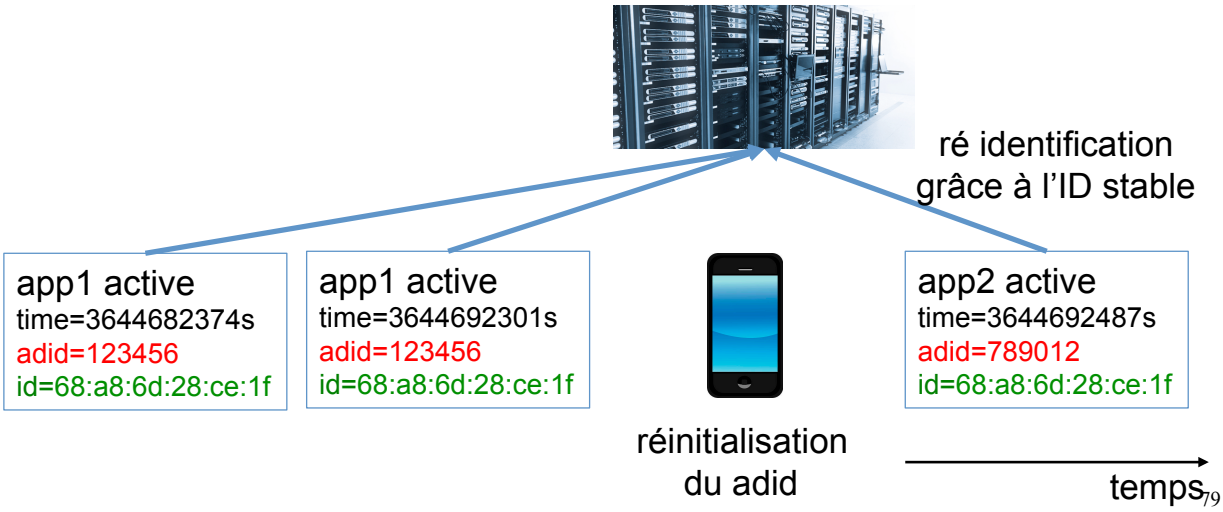
- **stable IDs** are perfect to **correlate** information collected from several Apps
  - ❍**and therefore create a user profile**

From the same device? It's sufficient to compare the IDs

If yes, we know a subset of Apps for this user and his/her centers of interest

About stable IDs and their use… (cont.')

- **stable IDs** are perfect to **bypass** the desired limits of advertising tracking
  - **if the user resets his Advertising ID, the A&A company can easily re-identify the user**



A GPS in your pocket… for others

To know more… (in French)

Retrouvez-nous sur notre site [www.cnil.fr/ip] en flashant le code ou sur :

Mobilitics, saison 2 :
Les smartphones et leurs apps sous le microscope de la CNIL et d'Inria

La CNIL et Inria travaillent depuis maintenant 3 ans sur un projet de recherche et d'innovation ambitieux nommé Mobilitics. Son objectif : mieux connaître les smartphones, ces objets utilisés quotidiennement par des dizaines de millions de français et qui

http://www.cnil.fr/fileadmin/documents/La_CNIL/publications/DEIP/Lettre_IP_N-8-Mobilitics.pdf

An example: My Talking Tom

« My Talking Tom » **accesses**

"imei": 8,
"network_code": 6,
"wifi_mac": 5,
"android_id": 12,
"operator_name": 8

« My Talking Tom » **transmits**

"android_id":
"85.195.69.168:(plain-text)",
"162.217.102.42:(plain-text)",
"vungle.com:(plain-text)",
"sponsorpay.com:(plain-text)"
"imei":
"ws.tapjoyads.com:(SSL)",
"1e100.net:(plain-text)",
"85.195.69.168:(plain-text)",
"outfit7.com:(plain-text)",
"sponsorpay.com:(plain-text)"
"wifi_mac":
"85.195.69.168:(plain-text)",
"vungle.com:(plain-text)",
"sponsorpay.com:(plain-text)"

Outline

1. introduction
   - ❍ two examples
   - ❍ "personal information" and the French/EU law
2. smartphones and personal information eco-system
   - ❍ why are we here?
   - ❍ let's come back to smartphones
   - ❍ who does what, who earns what?
   - ❍ where is the problem?
3. the manufacturer approaches for privacy parameter control
   - ❍ multiple limits
   - ❍ three quick examples
4. what we learned with the Mobilitics project
   - ❍ a rush towards stable identifiers… for a permanent user tracking that resists to resets
   - ❍ a GPS in your pocket… for the others
5. conclusions

Conclusions

mobile phone operators

+

smartphone manufacturers

The case of Google

Google

- **Google business model relies on advertisements**
- …and Google needs PI for that
  - ❍**Apps have an easy access to (stable) identifiers needed to track users**
    - sometimes without having to ask user authorization
  - ❍**very limited motivation to change the situation**
    - since August 2014, new Apps are supposed to only use the "Advertising ID" for targeted advertising…
    - … but it will take time and other identifiers still remain
    - current strategy remains to collect as many IDs as possible
  - ❍**and contrary indicators exist**
    - Android 4.3 proposed a privacy dashboard… Removed from the following Android versions!

The case of Google… (cont.')

- but this is (partially) an open-source OS
  - building secure versions is possible ☺
    - **BlackPhone (de Silent Circle)** 600 $
      - https://silentcircle.com/services#blackphone

    - **CryptoPhone 500 (de GSMK)** 3500 $
      - http://www.cryptophone.de/en/products/mobile/cp500/
      - can identify faked cell towers
        - http://www.popsci.com/article/technology/mysterious-phony-cell-towers-could-be-intercepting-your-calls
        - http://www.aftenposten.no/nyheter/iriks/Secret-surveillance-of-Norways-leaders-detected-7825278.html
      - usually those are "IMSI catchers"

The case of Apple

- **Apple sells (costly) hardware and softwares**
- … and communicates a lot on privacy

Tim Cook, PDG Apple : « Notre activité ne repose pas sur le fait de détenir des informations sur vous. Vous n'êtes pas notre produit »

  - **even if the situation is not perfect, there is are clear improvements across iOS versions**
    - many stable identifiers have been removed from the latest iOS versions
    - the Advertising-ID that a user can re-initialize is a key to limit tracking

The user can also

- limit the number of Apps
  - **be careful W.R.T. the App permissions asked or the privacy control dashboard**
  - **… and remove unused Apps**
  - **think it twice before using a daily assistant like "Google Now"**
- use official App stores
  - **Apps are checked (up to a certain point) by the store owner**
- switch off the Wifi interface if not used…
  - **to avoid physical tracking by stores (and others)**
- …and if you can, switch off data communications
  - **when not used**

The user can also… (cont.')

- explicitly stop Apps
  - **instead of leaving them running in background**
- set appropriate geolocation parameters
- limit advertising tracking / reset the AdvertisingID
  - **but it will be more efficient with iOS than it is with Android**
- "last but not least", do not jaibreak/root your phone
  - **otherwise any App has a full access to smartphone**

Fortunately the regulator has a real power

- the **EU laws** continue to evolve in the right direction
  - ❍**new EU regulation on data protection**
  - ❍**true impacts on companies**
  - ❍**EU data protection agencies (e.g., CNIL in France) discuss in the G29 group**

"On the Internet, nobody knows you're a dog."

In 1993…

In 2015…

"Remember when, on the Internet, nobody knew who you were?"